

СОЦИАЛЬНО-ЭКОНОМИЧЕСКОЕ ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ С УЧЕТОМ ЭКСПЕРТНОЙ ИНФОРМАЦИИ

И.И. Исмагилов,

Казанский (Приволжский) федеральный университет, г. Казань

Ключевые слова: *социально-экономическое прогнозирование, агрегирование прогнозов, регрессионная модель, экспертные методы, оценивание регрессионной модели с ограничениями, временные ряды, трендовые модели.*

Прогнозирование широкого круга процессов различной природы является крайне важной задачей при принятии решений во многих сферах человеческой деятельности. При этом следует отметить особую значимость прогнозирования социально-экономических процессов.

Прогнозирование – это один из ключевых моментов при принятии управленческих решений в социально-экономических системах. Регулярное прогнозирование процессов систем этого класса позволяет не только принимать эффективные управленческие решения, но и накапливать опыт, позволяющий повысить точность и надёжность прогнозов, улучшить модели и алгоритмы прогнозирования [1].

Выбор конкретного метода является одной из наиболее важных задач прогнозирования и связан с возрастанием актуальности разработки формальных, в том числе логических процедур. При этом можно указать три основные группы причин, влияющих на необходимость выбора метода прогнозирования [2]:

- увеличение числа методов прогнозирования, которое и в перспективе будет возрастать в связи с возрастанием практических задач прогнозирования (в настоящее время насчитывается уже несколько сотен методов прогнозирования);
- существенное возрастание сложности, как самих задач, так и объектов прогнозирования (создание корпоративных групп, холдингов, объединений и других сложных организационно-производственных структур).
- возрастание динамичности (подвижности) рыночной среды, ускорение темпов инновационного процесса.

Условно все существующие методы прогнозирования можно разбить на две большие группы [2]: фактографические (формализованные) и экспертные (интуитивные). Зачастую к вышеперечисленным добавляют еще одну группу – комбинированные (гибридные) методы, которые представляют собой сочетание методов этих двух групп прогнозирования.

Фактографические методы базируются на фактически имеющейся информации об объекте прогнозирования. В основном информационной базой для этих методов являются статистические данные.

Экспертные методы используют мнения экспертов (специалистов в соответствующей области), опирающихся на их знания, интуицию и опыт. Они применяются тогда, когда невозможно формализовать изучаемые процессы или имеет место неопределенность развития социально-экономической системы. Следует отметить, что экспертное прогнозирование в настоящее время единственный инструмент выработки прогнозов в условиях слабой определенности и структурированности объектов прогнозирования, особенно при долгосрочном периоде упреждения прогноза.

Гибридные методы, одновременно учитывающие в моделях прогнозирования статистическую и экспертную информацию [3,4]. Использование такой разнородной информации может повысить точность прогнозов. При этом в гибридных методах, кроме статистических данных, часто используются суждений экспертов об ожидаемом поведении изучаемого процесса на прогнозном периоде.

Следует также отметить, что в последнее время получили большое развитие методы прогнозирования с помощью гибридных систем, в основе которых лежат нечеткая логика, нейронные сети, генетические алгоритмы и другие новые информационные технологии [5,6]. Развиваются также методы нечеткого прогнозирования временных рядов на основе процедур коллективной экспертизы [7–9].

В данной работе предлагается метод прогнозирования на основе регрессионного моделирования с учетом экспертных суждений об интервалах возможных значений прогнозируемого показателя на периодах упреждения прогноза. Прогноз вырабатывается агрегированием ряда частных прогнозов, полученных с использованием регрессионных моделей с линейными ограничениями на параметры в виде равенств.

Краткое поэтапное описание метода прогнозирования имеет следующий вид.

1. Экспертное оценивание интервалов возможных значений прогнозируемого показателя на периодах упреждения прогноза ($i = 1, 2, \dots, p$):

$$[y_{i*}, y_i^*], \quad i = 1, 2, \dots, p, \quad p < k,$$

где k – количество регрессоров модели.

2. Формирование дискретных сеток значений прогнозируемого показателя на периодах упреждения размерности m (выбор за исследователем):

$$y_{ij} = y_{i*} + \Delta_i l, \quad l = 1, 2, \dots, m, \quad \Delta_i = \frac{y_i^* - y_{i*}}{m-1}.$$

3. Формирование линейных ограничений в виде равенств для параметров регрессионной модели на основе прогнозных значений регрессоров и определенного набора значений дискретных сеток для прогнозируемого показателя:

$$R \cdot b = r,$$

где b – вектор-столбец оценок параметров регрессионной модели с ограничениями (оценки условного МНК) размерности $(k+1)$; r – известный вектор-столбец, состоящий из p элементов (значения из дискретных сеток); R –

известная матрица порядка $p \times (k+1)$ (элементы представляют прогнозные значения регрессоров).

4. Определение вектора оценок параметров \mathbf{b} модели с ограничениями на параметры на основе вектора-столбца регрессанта \mathbf{Y} , минимизирующий сумму квадратов ошибок регрессии (остатков) с использованием метода множителей Лагранжа[9]:

$$\mathbf{b} = \hat{\mathbf{b}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\mathbf{b}}),$$

где \mathbf{X} регрессионная матрица; $\hat{\mathbf{b}}$ вектор-столбец параметров модели без ограничений (безусловные МНК-оценки параметров), определяемый по формуле

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

5. Вычисление регрессионной статистики и показателей предсказывающих свойств.

6. Выполнение этапов 3–5 до полного перебора комбинаций значений дискретных сеток для прогнозных значений регрессанта.

7. Выбор подмножества лучших моделей заданной размерности на основе выбранного критерия.

8. Агрегирование прогнозов лучших моделей и выработка окончательного прогноза.

Следует отметить, что отдельные этапы метода могут иметь различные алгоритмические реализации в зависимости от выбранных критериев отбора лучших моделей и метода агрегирования частных прогнозов.

Тестовую апробацию предлагаемого метода прогнозирования проведем решением задачи выработки ретроспективного прогноза (ex-post прогноза) налоговых поступлений в бюджет Забайкальского края на 2009 г. по статистическим данным с 2000 по 2008 гг. [11]. Прогнозирование в этой работе проводилось с использованием трендовых моделей (регрессий временного ряда налоговых поступлений (Y) на переменную условное время t). Были построены ex-post прогнозы по моделям без ограничений и с интервальным ограничением на прогнозные значения. Итоговое решение 3 экспертов: всего налоговых доходов в бюджет края в 2009 г. поступит от 16500 до 18000 млн. руб. Оценивание двух вариантов трендовых моделей проводилось безусловным МНК и решением задачи линейного программирования. Точность ex-post прогнозов оценивалась абсолютной относительной ошибкой прогноза на 1 лет вперед APE в %. Полученные результаты приведены в табл.1.

Таблица 1

Функция тренда	APE (модель без ограничений)	APE (модель с ограничениями)
Линейная	0,3	1,2
Квадратичная	0,4	1,2
Кубическая	22,4	7,7

Экспонента	32	5,5
Кривая Гомперца	11,9	7
Логистическая кривая	4,4	1,2

На основе проведенных исследований сделан вывод, что для прогнозирования общего количества налоговых поступлений на 2009 г. по однородной информации (статистическим данным) можно использовать линейный тренд вида $\hat{y}_t = 1372,13 + 1538,47t$, $R^2 = 0,94$. Модель статистически значима на уровне значимости 0,01. Прогноз на 2009 г. 16757, фактическое значение было 16703,3 млн. руб. ($APE=0,3\%$).

Модель линейного тренда по разнородной информации имеет вид $\hat{y}_t = 1026,88 + 1547,32t$. Прогноз на 2009 г. по ней 16757 млн. руб. ($APE=1,2\%$).

Рассмотрим результаты задачи прогнозирования [11] с использованием предлагаемого метода на основе модели линейного тренда. Метод реализован в виде скрипта в среде эконометрического пакета Gretl. Отметим, наличие лишь одного ограничения на параметры в виде некоторые матрицы в векторно-матричном соотношении для оценок параметров сводятся к скалярным величинам. При реализации было использовано разделение возможного интервала прогнозного значения показателя на 9 подинтервалов (равномерная дискретная сетка из 10 значений). Размерность подмножества лучших моделей выбралось на уровне 3, критерий отбора остаточная сумма квадратов отклонений RSS . Средняя абсолютная процентная ошибка $MAPE$ вычислялась по 4 последним наблюдениям обучающей выборки.

Полученные лучшие модели линейного тренда следующие:

1) $\hat{y}_t = 1483,98 + 1506,97t$, $RSS = 8655923$, $MAPE = 0,9\%$;

2) $\hat{y}_t = 1392,43 + 1532,79t$, $RSS = 85827,7$, $MAPE = 0,1\%$;

3) $\hat{y}_t = 1300,88 + 1538,47t$, $RSS = 861225,4$, $MAPE = 1,1\%$.

Модель линейного тренда без ограничений имела следующие показатели: $RSS = 8580178$, $MAPE = 2,1\%$.

Агрегирование частных прогнозов 3 лучших моделей в коллективный прогноз проводилось с использованием их взвешенной суммы. При этом веса вычислялись двумя способами: обратно пропорциональными RSS и $MAPE$. Получены следующие коллективные прогнозы:

1) веса, пропорциональны RSS : 16720,05 млн. руб., $APE=0,1\%$;

2) веса, пропорциональны $MAPE$: 16725,32 млн. руб., $APE=0,13\%$;

Прогноз по модели линейного тренда без ограничений был следующим: 16567 млн. руб., $APE=0,32\%$. Как видно, лучший коллективный прогноз обеспечивает более высокую точность.

В заключение отметим перспективность использования предложенного метода прогнозирования при решении широкого круга прогнозных задач, как на региональном уровне, так и на уровне предприятий. Например, среди них задачи прогнозирования показателей развития социально-экономической сферы региона [12, 13]. Обратим внимание также на расширение спектра задач

прогнозирования, связанного в настоящее время появлением новых аспектов социально-экономического развития в связи со становлением сетевой (цифровой) экономики [14]. При решении соответствующих задач прогнозирования наблюдается возрастание уровня информационной неопределённости относительно изучаемых процессов. Это связано с тем, что в сетевой экономике может наблюдаться также известное предположение П. Дэвида, получившее название “гипотеза отсрочки Дэвида” (David delay hypothesis) и гласящее о том, что происходит откладывание во времени выгод от применения информационных технологий [15].

Метод прогнозирования основан на использовании разнородной информации и позволяет получить более обоснованные прогнозы при адекватности экспертной информации динамике развития изучаемого социально-экономического процесса на периоде упреждения. В случае использования в методе регрессий на временных рядах и полиномиальных трендов для получения прогнозных значений факторов эффективное применение могут найти алгоритмы их оценивания на основе дискретных преобразований [16–18]. Относительная простота алгоритмических реализаций метода позволяет создавать простые программные средства для решения задач прогнозирования социально-экономических процессов на основе одновременного учета статистической и экспертной информации.

Список литературы

1. Жариков О.Н., Королевская В.И., Хохлов С.Н. Системный подход к управлению / Под ред. В. А. Персианова. М.: ЮНИТИ ДАНА, 2001. 62 с.
2. Глуценко В.В. Менеджмент: системные основы. Издание 2-е. М.; НПЦ – «Крылья», 1998. 224 с.
3. Головченко В.Б. Прогнозирование временных рядов по разнородной информации. – Новосибирск: Наука, 1999. 88 с.
4. Головченко, В.Б. Прогнозирование с использованием разнородной информации. – Иркутск: Изд-во БГУЭП, 2005. 71 с.
5. Ярушкина Н.Г. Основы теории нечетких и гибридных систем: учебное пособие. М.: Финансы и статистика, 2004.
6. Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов: Учебное пособие. Ульяновск: УлГТУ, 2010. 320 с.
7. Исмагилов И.И., Зинкин В.А. Нечеткое прогнозирование количественных показателей сложных систем // Исследования по информатике. 2007. № 11. С. 49–56.
8. Исмагилов И.И., Бичурин Р.В. Нечеткие прогнозы: классификация и метод их разработки на основе процедуры групповой экспертизы // Фундаментальные исследования. 2014. № 11 (часть 6). С. 1240–1247.
9. Ismagilov, I.I., Khasanova, S.F., 2015. Short-Term Fuzzy Forecasting of Brent Oil Prices. Asian Social Science. Vol. 11(11). P. 60–67.
10. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика: учебник. М.: Экзамен, 2003.

11. *Черхарова Н.И.* Оценивание параметров при подборе уравнений трендов для прогнозирования налоговых поступлений в региональный бюджет по статистической и экспертной информации // Вопросы современной науки и практики. Университет им. В.И. Вернадского. 2010. № 7-9. С. 58–65.
12. *Аксянова А.В., Хайрутдинова Ю.В.* Прогнозирование показателей развития социально-экономической сферы региона // Вестник Казанского технологического университета. 2011. № 20. С. 305–310.
13. *Kadochnikova E.I., Ismigilov I.I.* Measurement of cognitive growth factors of regional economy based on panel data // Mediterranean Journal of Social Sciences. 2014. Vol. 5(24). P. 377–382.
14. *Исмагилов И.И.* Стратегическое управление развитием предприятия в условиях становления сетевой экономики // Казанский экономический вестник. 2012. № 1. С. 16–18.
15. *Исмагилов И.И., Белов А.И.* Методические аспекты выбора портфеля проектов интеграции корпоративной информационной системы с инструментами электронной коммерции // Вестник КГФЭИ. 2010. № 4 (21). С. 64–69.
16. *Исмагилов И.И.* Дискретные преобразования в базисах уолш-подобных функций: Основы теории и применения в цифровой обработке сигналов / И.И. Исмагилов. Казань, Отечество, 2003. 130 с.
17. *Исмагилов И.И., Аглиуллин И.Н., Кирпичников А.П., Костромин А.В.* Полиномиальные модели трендов цифровых сигналов: алгоритмы оценивания на основе дискретных преобразований и сравнительный анализ их вычислительной сложности // Вестник Казанского технологического университета. 2015. Т. 18. № 12. С. 132–138.
18. *Ismagilov I.I., Khasanova S. F.* Algorithms of parametric estimation of polynomial trend models of time series on discrete transforms // Academy of Strategic Management Journal. 2016. Vol. 15. Special Issue. P. 21–28.